A Visual Odometry Method Based on the SwissRanger SR4000

Cang Ye* and Michael Bruch†

* University of Arkansas at Little Rock, 2801 S. University Ave, Little Rock, AR, USA 72204
† Space and Naval Warfare Systems Center Pacific, 53560 Hull Street, San Diego, CA 92152

ABSTRACT

This paper presents a pose estimation method based on a 3D camera—the SwissRanger SR4000. The proposed method estimates the camera's ego-motion by using intensity and range data produced by the camera. It detects the SIFT (Scale-Invariant Feature Transform) features in one intensity image and match them to that in the next intensity image. The resulting 3D data point pairs are used to compute the least-square rotation and translation matrices, from which the attitude and position changes between the two image frames are determined. The method uses feature descriptors to perform feature matching. It works well with large image motion between two frames without the need of spatial correlation search. Due to the SR4000's consistent accuracy in depth measurement, the proposed method may achieve a better pose estimation accuracy than a stereovision-based approach. Another advantage of the proposed method is that the range data of the SR4000 is complete and therefore can be used for obstacle avoidance/negotiation. This makes it possible to navigate a mobile robot by using a single perception sensor. In this paper, we will validate the idea of the pose estimation method and characterize the method's pose estimation performance.

Keywords: Visual odometry, pose estimation, SIFT feature detector, Flash LADAR.

1. INTRODUCTION

To perform autonomous navigation, a mobile robot has to maintain the awareness of its position. In case of no navigational infrastructure such as GPS, a robot may obtain position information through estimating the change of its pose over time. For navigation in a 3D environment, a robot pose includes its attitude (roll, pitch, yaw) and position (X, Y, Z coordinates). The existing robot pose estimation methods can be classified into two categories: proprioceptive and exteroceptive approaches. A typical proprioceptive method is to integrate the measurements from an Inertial Measurement Unit (IMU) and wheel odometry [1] [2]. The measurement accuracy of an IMU is subject to bias drift that accrues errors in attitudes over time. Wheel odometry may not provide accurate position estimates due to wheel slip. The inertial measurement method may fail when excessive wheel slip occurs (e.g., when a robot moves on loose soil). The representative exteroceptive approach in the literature is the Visual Odometry [3] (VO). The VO method employs a stereovision system to estimate the ego-motion by detecting features in a stereo image pair and tracking them from one frame to the next. The features' 3D positions in each frame are determined by stereo matching. Feature tracking between frames is performed by selecting features in an image and locating them in the subsequent image using spatial correlation search. The search method [3] [4] [5] requires an initial estimate of the motion that is obtained from wheel odometry and inertial sensor. Recently, a different type of feature matching approaches [6] [7] has been employed for VO. These approaches select and match features using the descriptor of each feature. They require stable and salient feature descriptors to work well with large image motions. But the advantage is that they do not perform spatial correlation search and thus no initial motion estimate is required for feature matching. The approach in [6] assumes very small motions between image frames. This requires the VO to be run in a high frequency and thus incurs a high computational cost. The approach in [7] uses a dense stereo depth map to compute the spatial invariants of features for feature matching. It is computationally expensive to generate a dense depth map. The use of a stereovision imposes other limitations on autonomous navigation. First, the depth accuracy of a stereovision system drops quadratically with the true distance. This means that only the data within a short range can be effectively used for the VO. Second, a stereovision system can not produce a complete set of range (depth) data of the environment in its field of view. As a result, stereovision data is not reliable for obstacle avoidance/negotiation. This problem is usually resolved by using additional perception sensor(s) such as a LADAR. This multi-sensor approach is not suitable for navigating a small robot that has strict size, weight, power and fidelity requirements for its onboard sensor [8].

Unmanned Systems Technology XII, edited by Grant R. Gerhart, Douglas W. Gage, Charles M. Shoemaker,
Proc. of SPIE Vol. 7692, 76921I · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.850349

The ideal solution to the autonomy of small robots (including portable/wearable robotic devices) is to use a single sensor modality for both pose estimation and obstacle avoidance/negotiation. The recent advance in Flash LADAR has created opportunity to address this autonomy issue. The Flash LADAR is a new class of 3D imaging sensor that can simultaneously generate intensity and range images of its environment. It is able to produce tens of image frames per second. The sensor has the following advantages over a stereovision system: (1) it measures depth by Time-of-Flight and therefore has consistent measurement accuracy in its full range; and (2) it returns a complete set of depth data of its environment. These advantages make a Flash LADAR suitable for the single-sensor-modality approach. In this paper, we propose and validate a pose estimation method based on a Flash LADAR—the SwissRanger SR4000. The method selects and matches features based on their descriptors. It works well with a large motion between image frames without the need of initial motion estimate.

The remainder of the paper is organized as follows: In Section II the SR4000's specification and the motivation of using the sensor is introduced. In Section III the technical details of the pose estimation method is presented. Then, in Section IV, the proposed method is validated and its pose estimation performance is characterized through experiments. The paper is concluded in Section V.

2. THE SWISSRANGER SR4000

The SwissRanger SR4000 (Fig. 1) is a Time-of-Flight (TOF) 3D camera developed and manufactured by MESA Imaging [9]. The TOF is determined by phase shift measurement. The camera illuminates its environment by amplitude modulated infrared light (850 nm). Its CMOS imaging sensor measures the phase shift of the returned modulated signal at each pixel. The amplitude and phase shift of the signal are used to determine the intensity and distance of the target point. As a result, the camera delivers both intensity and range image for each captured frame. The camera has a spatial resolution of 176×144 pixels and a field of view of 43.6×34.6 degrees. Its non-ambiguity range is 5 meters when a modulation frequency of 30 MHz is used. The programmable modulation frequency (29/30/31 MHz) allows simultaneous measurements of three cameras without interference. The camera has a small physical dimension: 50×48×65 mm^3. This makes it an ideal sensor for small-sized robots. Table 1 shows the sensor's main specifications.

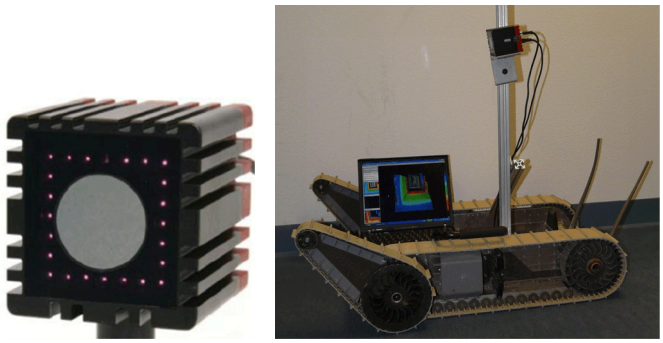


Fig. 1. The SR4000 and the Packbot robot with the sensor installed

Table 1 Specification of the SR4000 (Model: SR 00400001)	
Non-ambiguity range	5 m
Absolute accuracy	±10 mm (typ.)
Repeatability (1 σ)	4 mm (typ.), 7 mm (max.)
Pixel array size	176 (h) x 144 (v)
Field of view	43.6° (h) x 34.6° (v)
Modulation frequency	29/30/31 MHz
Acquisition mode	Continuous/ triggered
Integration time	0.3 to 25.8 ms
Dimension	65 x 65 x 68 mm
Weight	470

typ.: typical, max.: maximum, h: horizontal, v: vertical

Compared with a stereovision system, the SR4000 has larger power consumption since it uses light for illumination. However, this disadvantage is outweighed by the following advantages in navigating a mobile robot: First, the SR4000 measures depth based on phase shift and it has a consistent measurement error (±1 cm) for a distance up to 5 meters; while a stereovision system determines depth by computing disparity of the left and right images and the measurement

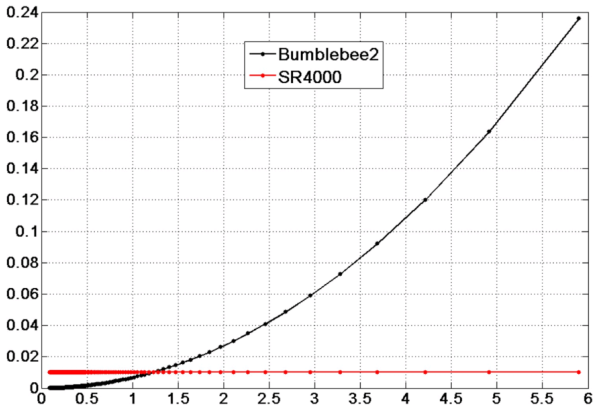
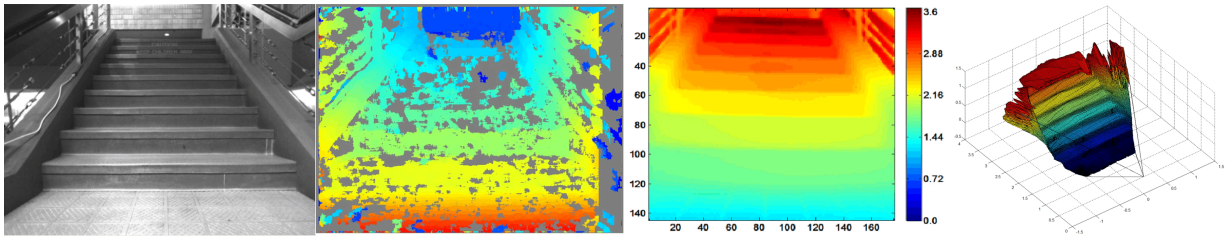


Fig. 2. Range measurement error of the SR4000 and Bumblebee2 stereovision system: the SR4000 (red) has a consistent error of 1 cm for a range up to 5 m while the stereovision's error (black) increases quadratically with range (~1 cm at 1.25 meters, and ~17 cm at 5 meters).

accuracy drops quadratically with true distance. Figure 2 compares the depth measurement accuracy of the SR 4000 and the Bumblebee2 stereovision system. The usefulness of the stereovision system's data for navigation is within ~2 meter, a much shorter range compared with the 5-meter range of the SR4000. The consistent measurement accuracy of the SR4000 indicates that a VO method based on the SR4000 may achieve much accurate pose estimation than a stereovision-based approach. Second, the SR4000's range data is complete and thus can be used for obstacle avoidance/negotiation. In contrast, a stereovision system's range data is incomplete and is not reliable for that purpose. Figure 3 compares the range images of the SwissRanger and the Bumblebee 2 stereovision system. The stereovision's range image contains large amount of missing data while the Swissranger's data is complete. The missing data makes it difficult to analyze the range data and use it for navigation. From the above discussion, we can see that the SR4000 is promising for the autonomy of small robots: a single perception sensor may be used for both pose estimation and obstacle avoidance/negotiation. In this paper, we will investigate the pose estimation problem. Our proposed pose estimation method uses both visual data (intensity image) and range data (range image) produced by the SR4000. To distinguish it from the VO method, we term our method VR-odometry.



(a) (b) (c) (d)

Fig. 3. Comparison of range images between the SwissRanger and the Bumblebee 2 system: (a) left image of the Bumblebee 2, (b) range image of the Bumblebee 2, (c) range image of the SR3000, (d) 3-D plot of the range data in (c).

3. VR-ODOMETRY

The idea of the VR-odometry method is to perform feature detection and matching in the SR4000's intensity images and use the matched features' range data to determine the pose change in two consecutive image frames. In this study, we employ a SIFT (Scale-Invariant Feature Transform) feature detector [10] to extract feature in an intensity image and match them to the SIFT features in the next intensity image. As the 3D coordinates of the matched features are known from the range data, the feature-matching process in fact solves the 3D points correspondence (data association) problem at each two sampling steps, and result in two 3-D data sets, $\{p_i\}$ and $\{p_i^{'}\}$; $i = 1,\cdots,N$. N is the number of matched SIFT features in the two image frames. The pose estimation problem is then to find a rotation and translation matrices, R and T, that minimize the error residual

$$e^2 = \sum_{i=1}^{i=N} \left\| p_i^{'} - Rp_i - T \right\|^2 . \qquad (1)$$

This least-square data sets fitting problem can be solved by the Singular Value Decomposition (SVD) method in [11]. As feature-matching in intensity images may result in incorrect data association (outliers), a RANSAC (Random Sample Consensus) process is implemented to reject the outliers. The entire method is as follows:

1) Extract the SIFT features in two consecutive images, find the matched features, and locate the corresponding 3D data sets $\{p_i\}$ and $\{p_i^{'}\}$.

2) Randomly select 4 associated points from the two data sets and form $\{p_k\}$ and $\{p_k^{'}\}$; $k = 1,\cdots,4$. Then find the least-square rotation and translation matrices ($\hat{R}_k$ and $\hat{T}_k$) for $\{p_k\}$ and $\{p_k^{'}\}$.

3) Project the entire data set $\{p_i\}$ onto $\{p_i^{'}\}$ using the found transformation ($\hat{R}_k$ and $\hat{T}_k$) and compute the error

$$e_i^2 = \left\| p_i^{'} - \hat{R}_k p_i - \hat{T}_k \right\|^2 \qquad (2)$$

for each data-pair $i = 1,\cdots,N$. A threshold ε is used to score S_k for this transformation: S_k is incremented once for each $e_i^2 < \varepsilon$.

4) Step 2 and 3 repeat for a fixed number of iterations or until exhausting all combination of point set selections, whichever is smaller. The transformation with the highest score is recorded. The corresponding data sets $\{p_j\}$ and $\{p_j^{'}\}$; $j = 1,\cdots,S_k$, where each data-pair satisfy the threshold test in step 3, are selected and used to compute the maximum likelihood transformation estimate $\hat{R}$ and $\hat{T}$ by the SVD least-square fitting method. The robot's rotation can be computed from $\hat{R}$ and its translation is determined by $\hat{T}$.

3.1 Determination of the Least-Square Transformation

We use the following SVD algorithm to determine the least-square rotation and translation matrices $\hat{R}$ and $\hat{T}$ of two corresponding data sets $\{p_m\}$ and $\{p_m^{'}\}$; $m = 1,\cdots,M$ ($M \geq 4$):

(i) Compute the centroid p and $p^{'}$ of $\{p_m\}$ and $\{p_m^{'}\}$, and then calculate

$$q_m = p_m - p \quad \text{and} \quad q_m^{'} = p_m^{'} - p^{'}. \tag{3}$$

(ii) Compute the 3×3 matrix

$$\Omega = \sum_{m=1}^{M} q_m q_m^{'\,t}, \tag{4}$$

(iii) Find the SVD of Ω:

$$\Omega = U \Lambda V^t, \tag{5}$$

where U and V are 3×3 orthogonal matrices and Λ=diag(λ_1, λ_2, λ_3) is a diagonal matrix with non-negative elements. λ_1, λ_1, λ_2 are the singular values.

(iv) Calculate *det*(*Y*), the determinant of *Y*.

(a) If *det*(*Y*)=+1, then

$$\hat{R} = VU^t \tag{6}$$

(b) If *det*(*Y*)=-1 and one of the singular values equals zero, i.e., λ_l=0 (l=1, 2, or 3) , then

$$\hat{R} = V^{'}U^t, \tag{7}$$

where $V^{'}$ is obtained from V by changing the sign of the l^{th} column.

(c) Otherwise, no solution can be found by the approach. This may happen only if the data sets incur large noise. This is not likely the case as the SR4000's range error is about 10 mm with a repeatability of 7 mm. But if this happen, the current frame is skipped and the next frame is used for pose estimation.

(d) Once the solution to $\hat{R}$ is found, the translation matrix can be computed by

$$\hat{T} = p^{'} - \hat{R}p. \tag{8}$$

The proof of the above SVD method is referred to [11].

3.2 Determination of pose change

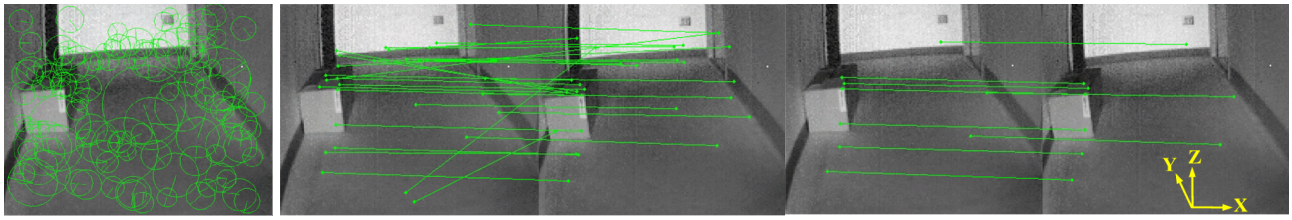
The change of attitude between the two image frames can be computed from the least-square rotation matrix, $\hat{R} = \{r_{ij}\}$ for $i = 1,2,3$ and $j = 1,2,3$, by

$$\begin{cases} \psi = a\tan 2(r_{22}, -r_{11}) \\ \theta = \text{atan2}(\text{r}_{32}, -\text{r}_{12}\sin\psi + \text{r}_{22}\cos\psi) \\ \phi = \text{atan2}(\text{r}_{13}\cos\psi + \text{r}_{23}\sin\psi, \text{r}_{11}\cos\psi + \text{r}21\sin\psi) \end{cases} \quad (9)$$

In this work, we use the Euler angles (roll ϕ, pitch θ, yaw ψ) to describe the attitude. The change in X, Y and X coordinates corresponds to the x, y, z components of the 3×1 matrix $\hat{T}$.

4. CHARACTERIZATION OF THE VR-ODOMETRY

A rudimentary implementation of the VR-odometry has been performed in Matlab environment to validate the idea. The implementation uses SIFT feature descriptors as they are invariant to translations, rotations and scaling and robust to residual small distortions. Figure 4 illustrates how the VR-odometry finds the correct correspondences between features in two successive images. The images were captured when a human handheld the SR4000 and moved in a hallway. Figure 4a depicts the detected SIFT features. They were matched with the features in the next frame based on the SIFT features' scales and orientations. Figure 4b shows the initial matches that exhibit some outliers (mismatches). The RANSAC process removed the outliers and the results are shown in Fig. 4c.



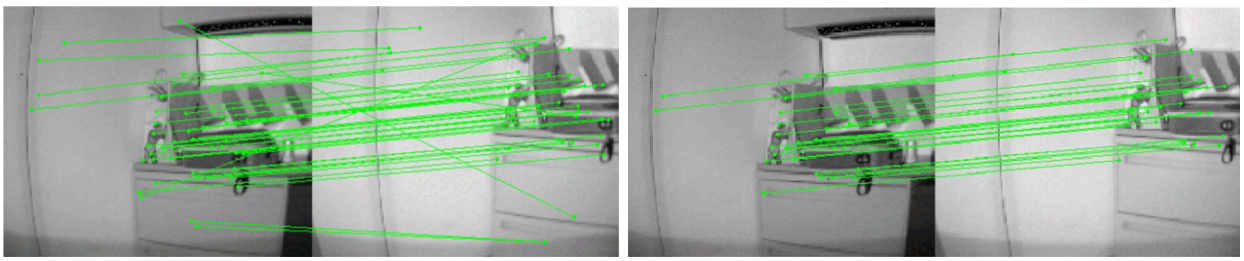
(a) (b) (c)

Fig. 4. SIFT feature matching and RANSAC for outlier removal: (a) SIFT features of an image frame (each circle and the straight line represent the scale and orientation of the SIFT feature located at the center of the circle.); (b) Initial matched features in two consecutive image frames; (c) Matched features after RANSAC. Intensity images were captured in a hallway. The time interval between the right and left images in (b) is 100 milliseconds. Therefore, the objects in the right image are closer to the sensor.

A series of experiments were carried out in an office environment with objects ranging 1.5-4.5 meters from the SR4000. The sensor was installed on a pan-tilt unit. Experimental runs, with various combinations of pitch, yaw rotation and X, Y translations, were performed to quantify the measurement accuracy and repeatability. In all experiments, roll ϕ and Z translation are always zero. 1000 images for each pose were taken to compute pose changes and the error statistics. It is noted that in all experiment we use raw sensor data for pose estimation.

4.1 Distribution of pose measurement errors

We carried out two experiments to examine the distribution of pose measurement errors. In the first experiment, the SR4000's pose change was zero. The experiment result demonstrates that the measurement error of pose (ϕ, θ, ψ, x, y, z) is zero-mean Gaussian with standard deviation (0.1°, 0.2°, 0.2°, 7 mm, 3 mm, 6 mm). The result indicates that the VR odometry's inherent error is a white Gaussian noise. Also, the measurement accuracy (mean error) and repeatability (standard deviation) is very good, meaning the sensor's noise has limited effect on the VR-odometry's performance in pose estimation.



(a) Initial matched features (b) Matched features after RANSAC

Fig. 5. SIFT features matching with rotation and translation of the SR4000. The right image in (a) / (b) was captured after the rotation and translation.

In the second experiment, the SR4000 had a combination of rotation (θ=-5.9°, ψ=5.0°) and translation (X=-80 mm, Y=130 mm). Figure 5 show the SIFT features matching in two consecutive frames. The distribution of the pose measurement error is depicted in Fig. 6. It follows a normal distribution whose mean and standard deviation are (-0.1°, 0.2°, -0.3°, 8 mm, -2 mm, 11 mm) and (0.5°, 0.4°, 0.4°, 13 mm, 5 mm, 11 mm), respectively. Compared with the first experiment, the results exhibit a bias in the mean error and relatively larger standard deviation.

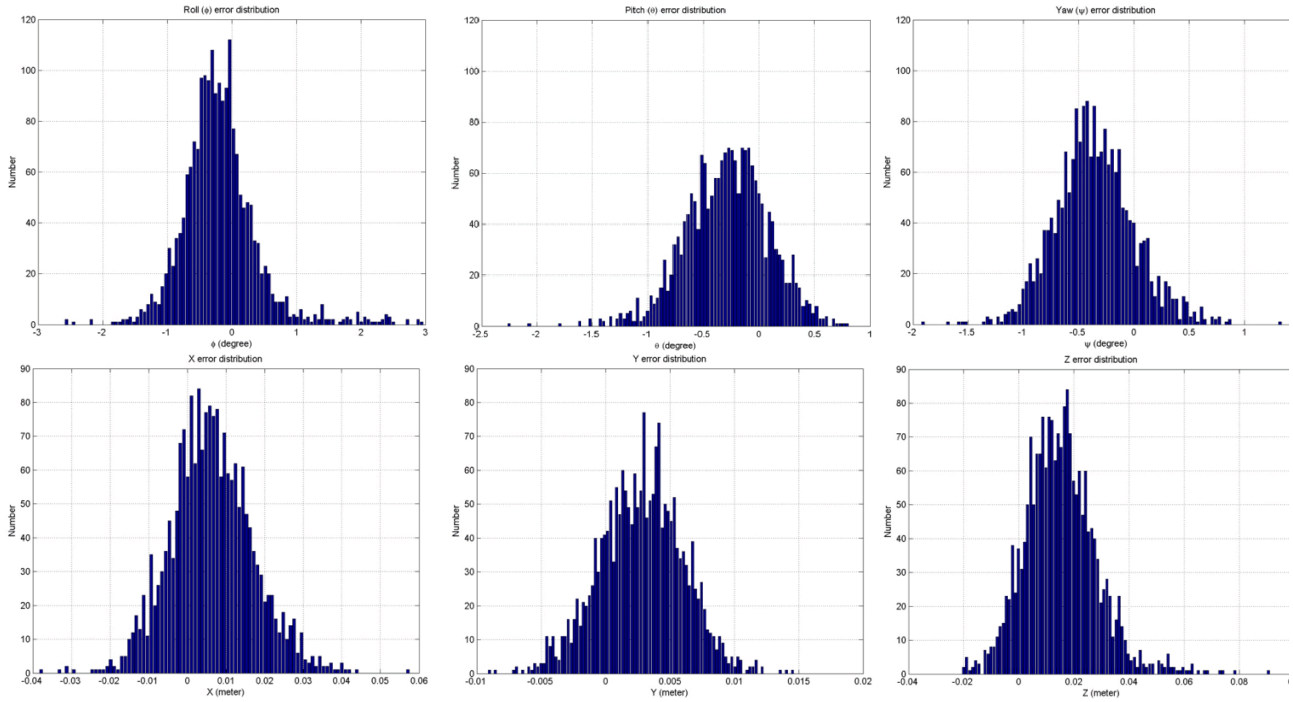
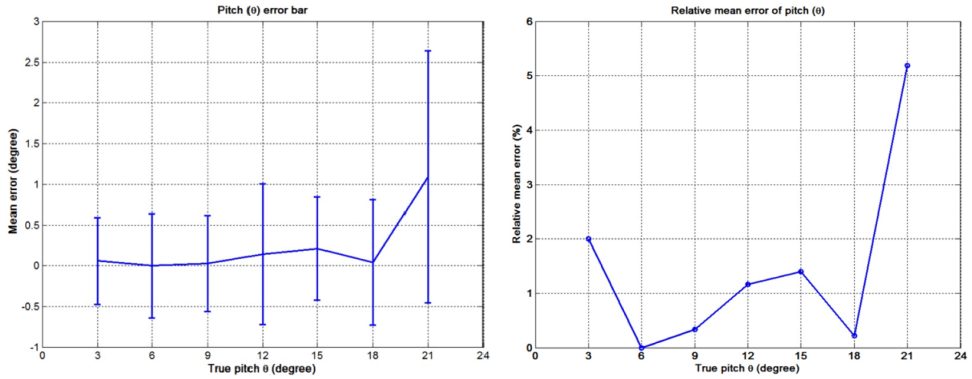


Fig. 6. Distribution of pose measurement error with rotation (θ=-5.9°, ψ=5.0°) and translation (X=-80 mm, Y=130 mm)

Both experiments show a decent accuracy in attitude measurement. Considering that the SR4000's average angular resolution is about 0.25°, the accuracy of attitude measurement is reasonable. The VR-odometry also demonstrates a very good accuracy and repeatability in measuring motion along Y axis due to the sensor's high accuracy and repeatability in depth measurement. The sensor's resolution in measuring motion along X and Z axis drops proportionally with the depth (Y) value. At Y=5 meters, the resolution is about 23 mm. The VR-odometry's measurement accuracy and repeatability in X and Z axes seem reasonable. However, a proper camera calibration and data filtering may improve the performance.



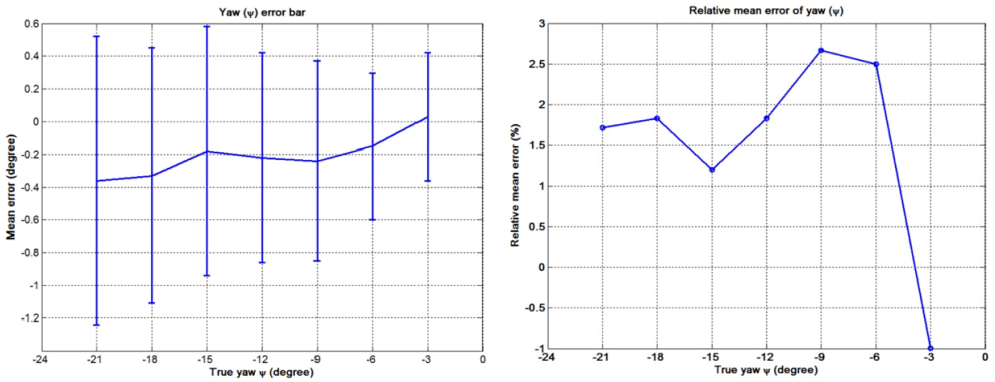
(a) Confidence (1σ) interval (b) Relative mean error

Fig. 7. Statistics of the pitch measurement (computed from 1000 samples)

4.2 Accuracy and repeatability of pose measurements

Experiments were performed to inspect the pose estimation performance for each individual motion. Due to the constraint of test facility, we have only done experiments with pitch rotation, yaw rotation, and translation along X and Y directions. In the first experiment, the sensor underwent a pitch rotation in the range [3°, 21°] (increment: 3°/step). 1000 images were captured before and after each pitch rotation for computing the pose change. Figure 7a and 7b depicts the error bars and the relative mean error of the pitch measurements, respectively. It can be observed that the accuracy (mean error) and the repeatability of the pitch measurement are quite decent for a pitch rotation in the range [3°, 18°]: the mean errors are in the range [0°, 0.21°] and the standard deviations are in [0.53°, 0.86°]. The relative mean errors are between 0.0% and 2.0%.

The second experiment is to test the VR-odometry's performance in measuring yaw rotation in the range [-21°, -3°]. The result is shown in Fig. 8. The mean errors are in the range [-0.36°, 0.03°] and the standard deviations are in [0.39°, 0.88°]. The relative mean errors are between -1.0% and 2.7%.



(a) Confidence (1σ) interval (b) Relative mean error

Fig. 8. Statistics of the yaw measurement (computed from 1000 samples)

Table 2 shows the statistics of the roll, pitch and yaw measurements of both experiments. It can be observed that the measurements are accurate if the pitch/yaw rotation is within ±18°, a large image motion between image frames compared to the camera's field of view. The mean errors are mostly within the camera's angular resolution (~0.25°). There are a few exceptions that the mean errors go beyond ±0.25° (the worst case: -0.46°). We believe that this can be improved if proper data filtering and sensor calibration will be performed. For measurement repeatability, the standard deviations of the roll measurements are consistently small. This is probably because that there was no roll rotation in the experiments. The repeatability of pitch and yaw measurements need to be improved in our future work.

Table 2 Measurement accuracy of rotation

MV: (μ, σ) / TV: (ϕ, θ, ψ)	Roll ϕ (°)	Pitch θ (°)	Yaw ψ (°)
(0, 3, 0)	(0.06, 0.27)	(3.06, 0.53)	(-0.16, 0.57)
(0, 6, 0)	(0.07, 0.16)	(6.00, 0.64)	(0.02, 0.66)
(0, 9, 0)	(0.01, 0.21)	(9.03, 0.59)	(0.24, 0.65)
(0, 12, 0)	(0.22, 0.21)	(12.14, 0.86)	(0.24, 0.69)
(0, 15, 0)	(0.11, 0.33)	(15.21, 0.63)	(-0.12, 0.91)
(0, 18, 0)	(0.00, 0.31)	(18.04, 0.77)	(-0.46, 0.80)
(0, 21, 0)	(0.29, 0.44)	(22.09, 1.54)	(0.14, 1.06)
(0, 0, -3)	(-0.03, 0.31)	(0.10, 0.52)	(-2.97, 0.39)
(0, 0, -6)	(-0.03, 0.19)	(-0.02, 0.50)	(-6.15, 0.45)
(0, 0, -9)	(0.02, 0.25)	(0.20, 0.60)	(-9.24, 0.61)
(0, 0, -12)	(-0.02, 0.26)	(0.15, 0.66)	(-12.22, 0.64)
(0, 0, -15)	(-0.09, 0.32)	(0.36, 0.68)	(-15.18, 0.76)
(0, 0, -18)	(-0.13, 0.30)	(0.15, 0.69)	(-18.33, 0.78)
(0, 0, -21)	(0.04, 0.39)	(0.20, 0.77)	(-21.36, 0.88)

MV: Measured Values, TV: True Values, μ: mean error, σ: standard deviation; Camera's angular resolution: pitch: 0.25º, yaw: 0.24º. 1000 samples were used to compute the statistics.

The third experiment is to test the measurements of translation when the sensor moves along X or Y directions. A step size of 305 mm was used in the experiments. We found that the VR-odometry did not produce satisfactory results if the translation is bigger than 610 mm. (This suggests that a smaller step size should be used in the future study.) The results are tabulated in Table 3, from which we can observe that the VR-odometry has highly consistent accuracy and repeatability in measuring movement along Y axis (depth). This is attributed to the sensor's consistent measurement accuracy in Y axis (±1 cm), an apparent advantage over the

stereovision-based approach. However, it has relatively larger mean errors and/or standard deviations in X/Z measurements (compared with the camera's resolution). These need to be improved in our future work.

The results in Tables 2 and 3 were computed using raw sensor data. Proper data filtering and camera calibration will improve the measurement accuracy and repeatability in the future work. It should be noted that in some cases the VR-odometry did not find a solution to the pose estimation. The error states were recorded and the data were discarded, meaning that the samples for computing some of the statistics in Tables 2 and 3 may be slightly smaller than 1000. Some of the failures were resulted because the RANSAC process did not find a sufficient number matched features (possibly due to the use of an overly small threshold ε). We have not yet looked into the cause of the other failure cases.

Table 3 Measurement accuracy of translation

MV: (μ, σ) / TV: (X, Y, Z)	X (MM)	Y (mm)	Z (mm)
(305, 0, 0)	(297, 16)	(-2, 7)	(-8, 24)
(610, 0, 0)	(579, 64)	(-9, 11)	(-4, 75)
(0, 305, 0)	(-15, 73)	(303, 12)	(35, 55)
(0, 610, 0)	(-0, 98)	(607, 17)	(83, 98)

MV: Measured Values, TV: True Values, μ: mean error, σ: standard deviation. Camera's resolution at 5 meters: X: 22 mm, Z: 23 mm, Y: 10 mm. 1000 samples were used to compute the statistics.

In our current implementation, we use the SIFT feature descriptors and the RANSAC method for the sake of the accuracy and reliability of pose estimation. Both approaches are computationally expensive. In our future work, we will develop a more efficient method with real-time performance. This may be achieved through the following efforts: (1) investigate other feature descriptors and adopt one with less computational cost such as the SURF (Speeded Up Robust Features) [12] for the VR-odometry; and (2) use spatial invariants between the detected features' 3D points (e.g., distances between the 3D points) for outlier rejection (or inlier detection) to remove or accelerate the RANSAC process. In term of improving the method's accuracy and repeatability, we will calibrate the camera, develop filtering method to filter out data with low confidence levels and develop feature sampling method to select features with more accurate depth values.

5. CONCLUSIONS

In this paper, we propose and validate a pose estimation method, termed VR-odometry. The method simultaneously uses the intensity and range data of a 3D camera—the SwissRanger SR4000—to determine the pose change of a robot. We have characterized the method's performance using the sensor's raw data and the results demonstrate that the VR-odometry is accurate in measuring attitude change and movement along Y axis and has high repeatability in the Y measurement. Further research will be carried out to improve the data repeatability of the measurements of attitude change, and the accuracy and repeatability in estimating translations along X and Z axes. The rudimentary implementation presented in this paper has demonstrated that the proposed method is promising for autonomous navigation of small-sized robots.

ACKNOWLEDGMENTS

This work was supported in part by a Summer Faculty Research Fellowship from the Office of Naval Research, a NASA EPSCOR RID grant, and by the Arkansas Space Grant Consortium under grant number SW1903. The experiments were carried out at the Space and Naval Warfare Systems Center Pacific.

REFERENCES

[1] G. Dissanayake, S. Sukkarieh, E. Nebot, and H. Durrant-Whyte, "The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 5, pp. 731-747, 2001.

[2] C. Ye, "Toward safe navigation in urban environments," *SPIE Proc. 6561: Unmanned Systems Technology IX*, Orlando, FL, April 9-12, 2007.

[3] Y. Cheng, M. W. Maimone, and L. Matthies, "Visual odometry on the Mars exploration rovers – a tool to ensure accurate driving and science imaging," *IEEE Robotics and Automation Magazine*, vol. 13, no. 2. pp. 54-62, 2006.

[4] M. Agrawal, K. Konolige, R. Bolles, "Localization and mapping for autonomous navigation in outdoor terrains: a stereo vision approach," *Proc. IEEE Workshop on Applications of Computer Vision*, 2007.

[5] P. Corke, D. Strelow and S. Singh, "Omnidirectional visual odometry for a planetary rover," *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 4007-4012.
[6] D. Nister, O. Naroditsky and J. Bergen, "Visual odometry," *Proc. IEEE Conference on Computer Vision Pattern Recognition*, 2004, pp.652-659.
[7] H. Hirschmüller, P. Innocent, and J. Garibaldi, "Fast, unconstrained camera motion estimation from stereo without tracking and robust statistics," *Proc. International Conference on Control, Automation, Robotics and Vision*, 2002, pp. 1099-1104.
[8] M. H. Bruch, *et al.*, "Advances in autonomy for small UGVs," *SPIE Proc. 5804: Unmanned Ground Vehicle Technology VII*, Orlando, FL, March 29-31, 2005.
[9] http://www.mesa-imaging.ch/prodview4k.php
[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 2, no. 60, pp. 91-110, 2004.
[11] K. S. Arun, et al, "Least square fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 698–700, 1987.
[12] H. Bay, *et al.*, "SURF: Speeded up Robust Features", *Proc. European Conference on Computer Vision*, 2006.